II.B Books, Papers, and Abstracts

The publications for the various collaborative projects are summarized in their respective progress reports.

II.C Resource Summary Table

Detailed resource usage information is summarized starting on page 36. Tabulation of this information on the NIH-provided form is yet to be completed.

Appendix A

Community Growth and Project Synopses

This appendix contains a graphical display of the development of the SUMEX-AIM community over the years and brief synopses of currently active projects. Figure 15 below illustrates the substantial growth in the cumulative number of projects in the Stanford, national SUMEX, and Rutgers-AIM communities since the resource began operation in 1974.

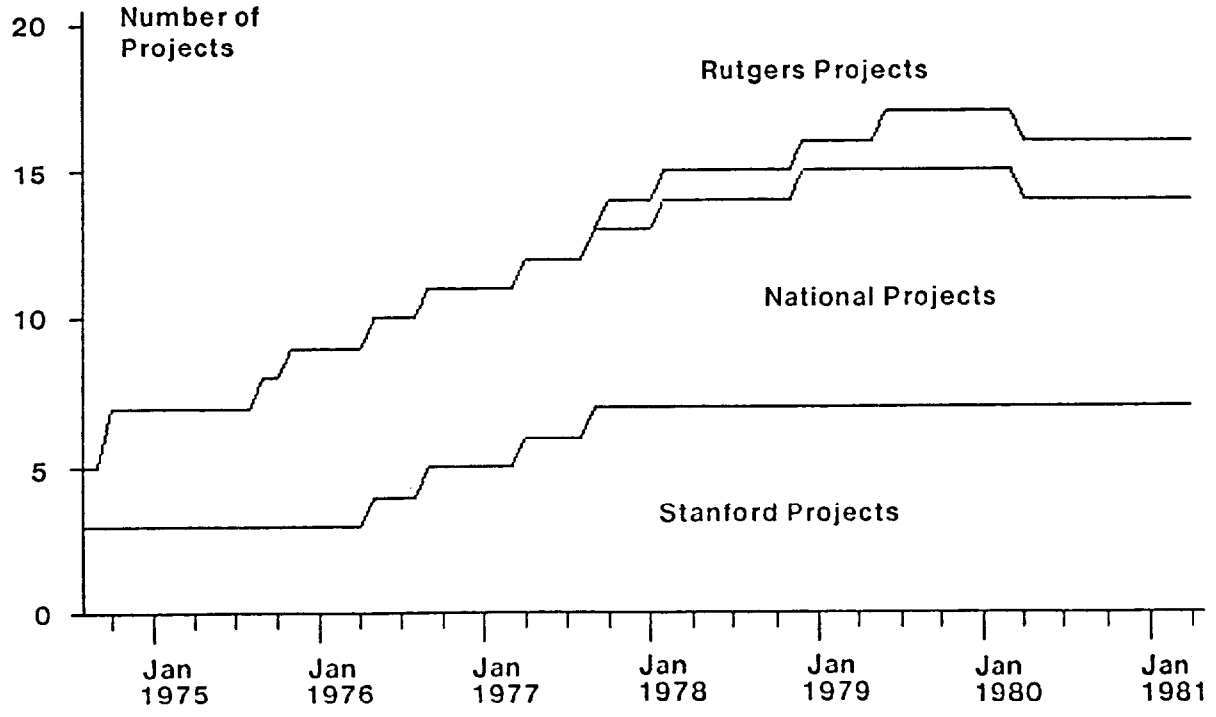

Figure 15. SUMEX-AIM Growth by Community

National AIM Project: ACQUISITION OF COGNITIVE PROCEDURES (ACT)

Principal Investigator: John R. Anderson, Ph.D.
Department of Psychology
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
(412) 578-2788 (ANDERSON@SUMEX-AIM)

The ACT Project combines a semantic network data-base with a production system to simulate human cognition. Prominent among the reasons for using a production system architecture as a framework for developing such a program is the possibility of modeling learning as the acquisition of new productions. ACT possesses a number of learning mechanisms which have been used to model the learning of procedural skills such as language comprehension and geometry theorem proving. Some of these mechanisms have the effect of either extending or restricting the set of circumstances in which a particular behavior is performed so as to produce better performance. Others have the effect of speeding up cognitive operations by compressing the effects of a series of production applications into the application of a single production. Out of this set of productions ACT applies those that usually result in desirable outcomes. In this way it is able to model the human ability to learn even when given unreliable feedback. Another feature of ACT that reflects its psychological orientation is its willingness to model human limitations. Here the hope is that by being faithful to the human mind even in its failings, it eventually may be possible to emulate its successes.

SOFTWARE AVAILABLE ON SUMEX

The ACT production system is available to GUEST users of SUMEX.

REFERENCES

Anderson, J.R.: Language, Memory, and Thought. Lawrence Erlbaum Associates, Hillsdale, N.J., 1976.

Anderson, J.R., Kline, P.J. and Lewis, C.H.: A production system model of language processing. IN M.A. Just and P.A. Carpenter (eds.), Cognitive Processes in Comprehension. Lawrence Erlbaum Associates, Hillsdale, N.J., 1977.

Anderson, J.R. and Kline, P.J.: A learning system and its psychological implications. Proc. Sixth IJCAI, Tokyo, August, 1979.

National AIM Project: CADUCEUS (previously INTERNIST)

Principal Investigators: Jack D. Myers, M.D. (MYERS@SUMEX-AIM)
Harry E. Pople, Ph.D. (POPLE@SUMEX-AIM)
University of Pittsburgh
Pittsburgh, Pennsylvania 15261
Dr. Pople: (412) 624-3490

The major goal of the CADUCEUS Project is to produce a reliable and adequately complete diagnostic consultative program in the field of internal medicine. Although this program is intended primarily to aid skilled internists in complicated medical problems, the program may have spin-off as a diagnostic and triage aid to physicians' assistants, rural health clinics, military medicine and space travel. In the design of CADUCEUS and its predecessor INTERNIST I, we have attempted to model the creative, problem-formulation aspect of the clinical reasoning process. The program employs a novel heuristic procedure that composes differential diagnoses, dynamically, on the basis of clinical evidence. During the course of a CADUCEUS or INTERNIST I consultation, it is not uncommon for a number of such conjectured problem foci to be proposed and investigated, with occasional major shifts taking place in the program's conceptualization of the task at hand.

SOFTWARE AVAILABLE ON SUMEX

Versions of INTERNIST are available for experimental use, but the project continues to be oriented primarily towards research and development; hence, a stable production version of the system is not yet available for general use.

REFERENCES

Pople, H.E., Myers, J.D. and Miller, R.A.: The DIALOG model of diagnostic logic and its use in internal medicine. Proc. Fourth IJCAI, Tbilisi, USSR, September, 1975.

Pople, H.E.: The formation of composite hypotheses in diagnostic problem solving: An exercise in synthetic reasoning. Proc. Fifth IJCAI, Boston, August, 1977.

National AIM Project: HIERARCHICAL MODELS OF HUMAN COGNITION

Principal Investigators: Walter Kintsch, Ph.D. (KINTSCH@SUMEX-AIM)
Peter G. Polson, Ph.D. (POLSON@SUMEX-AIM)
Computer Laboratory for Instruction
in Psychological Research (CLIPR)
Department of Psychology
University of Colorado
Boulder, Colorado 80302
(303) 492-6991

Contact: Dr. James Miller (JMILLER@SUMEX-AIM)

The CLIPR Project is concerned with the modeling of complex psychological processes. It is comprised of two research groups. The prose comprehension group has completed a project that carries out the microstructure text analysis described by Miller and Kintsch (1980), yielding predictions of the recall and readability of that text by human subjects. More recently, this group has been interacting with the Heuristic Programming Project at Stanford, using the AGE and UNITS packages to build a more complex model of the knowledge-based processes characteristic of prose comprehension. The planning group is working toward a model of the planning processes used by expert computer software designers. The initial development of this model requires the detailed analysis of expert software design protocols for subsequent simulation.

SOFTWARE AVAILABLE ON SUMEX

A set of programs has been developed to perform the microstructure text analysis described in Kintsch and van Dijk (Psychological Review, 1978) and Miller and Kintsch (1980). The program accepts a propositionalized text as input, and produces indices that can be used to estimate the text's recall and readability. A more complex model based in AGE and UNITS, which emphasizes the knowledge-based aspects of comprehension, is currently under development.

REFERENCES

Atwood, M.E., Polson, P.G., Jeffries, R. and Ramsey, H.R.: Planning as a process of synthesis. Technical Report SAI-78-144-DEN. Science Applications, Inc., Denver, Colorado, December, 1978.

Jeffries, R., Turner, A.A., Polson, P.G., and Atwood, M.A.: The processes involved in designing software. In J.R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, NJ: Erlbaum, 1981 (forthcoming).

Kintsch, W.: On modeling comprehension. Educ. Psychologist, 14:3-14, 1979.

Miller, J.R. and Kintsch, W.: Readability and recall of short prose passages: A theoretical analysis. J. Experimental Psychology: Human Learning and Memory, 1980. (In press)

National AIM Project: BIOMEDICAL KNOWLEDGE ENGINEERING IN CLINICAL MEDICINE (PUFF-VM)

Principal Investigators: John J. Osborn, M.D.
The Institutes of Medical Sciences
Pacific Medical Center
San Francisco, California 94115
(415) 567-0900 (OSBORN@SUMEX-AIM)

Edward A. Feigenbaum, Ph.D.
Department of Computer Science
Stanford University

The PUFF-VM Project has produced two knowledge-based programs for the interpretation of physiologic measurements made in clinical medicine. The interpretations are intended to aid in diagnostic decision-making and in selecting therapeutic actions. The programs are: PUFF--the evaluation of pulmonary function laboratory data, and VM--the evaluation and management of respiratory status for patients in the intensive care unit.

The task of the PUFF PROGRAM is to interpret standard measures of pulmonary function. In the laboratory at the Pacific Medical Center (PMC), about 50 parameters are calculated from measurement of lung volumes, flow rates, and diffusion capacity. In addition to these measurements, patient history and referral diagnosis also are used to interpret the test results. PUFF produces a report for the patient record, explaining the clinical significance of measured test results. It also provides a diagnosis of the presence and severity of pulmonary disease. The interpretation process is accomplished by examination of expert knowledge represented by a set of production rules. Each rule relates physiologic measurements or states to a conclusion about the physiologic significance of the measurement or state. A version of the PUFF program is used daily at the PMC.

The VENTILATOR MANAGER (VM) PROGRAM is designed to interpret on-line physiologic data in the intensive care unit (ICU). These data are used to manage post-surgical patients receiving mechanical assistance in breathing. VM is an extension of a physiologic monitoring system, and is designed to perform 5 specialized tasks in the ICU: 1) to detect possible measurement errors; 2) to recognize untoward events in the patient/machine system and suggest corrective action; 3) to summarize the patient's physiologic status; 4) to suggest adjustments to therapy based on the patient's status over time, and long-term therapeutic goals; and 5) to maintain a set of patient-specific expectations and goals for future evaluation by the program. The program produces interpretations of the physiologic measurements over time, using a model of the therapeutic procedures in the ICU and clinical knowledge about the diagnostic implications of the data.

SOFTWARE AVAILABLE ON SUMEX

The PUFF and VM programs will be available to GUEST users for use on pre-existing (non-identifiable) cases. No packages currently exist for program development.

REFERENCES

Fagan, L.M., Kunz, J.C., Feigenbaum, E.A. and Osborn, J.J.: A symbolic processing approach to measurement interpretation in the intensive care unit. Proc. Third Annual Symposium Computer Applications in Medical Care, Silver Spring, Maryland, October, 1979, pp. 30-33.

Fagan, L.M., Shortliffe, E.H. and Buchanan, B.G.: Computer-based medical decision making: From MYCIN to VM. Automedica 3(2), 1980.

Kunz, J.C., Fallat, R.J., McClung, D.H., et al: A physiological rule based system for interpreting pulmonary function test results. Heuristic Programming Project Report HPP-78-164, Computer Science Dept., Stanford Univ., November, 1978.

Osborn, J.J., Fagan, L.M., Fallat, R.J., et al: Managing the data from respiratory measurements. Med. Instrumentation, November-December, 1979.

Rutgers AIM Project: RUTGERS RESEARCH RESOURCE - COMPUTERS IN BIOMEDICINE

Principal Investigator: Saul Amarel, Ph.D.
Department of Computer Science
Rutgers University
New Brunswick, New Jersey 08903
(201) 932-3546 (AMAREL@RUTGERS)

The broad objective of the Resource is to apply advanced methods in computer science, particularly in artificial intelligence (AI), to biomedical problems. The Resource has three major areas of study: 1) Medical Modeling and Decision Making in several medical domains with emphasis on collaborative development of consultation systems in rheumatology and ophthalmology; 2) Modeling of Belief Systems and Commonsense Reasoning with emphasis on the psychology of plan recognition and handling of stereotypes; and 3) Artificial Intelligence studies with emphasis on Representations, Interpretation processes, and problems of knowledge and expertise acquisition. The studies in Medical Modeling and Decision Making are performed jointly by computer and medical scientists at Rutgers and elsewhere in the Country and abroad.

The Resource also sponsors national Artificial Intelligence in Medicine (AIM) Workshops for the AIM community.

SOFTWARE AVAILABLE ON SUMEX

CASNET--System for consultation in the diagnosis and treatment of glaucoma.

Documentation available:

"A Model-Based Method for Computer-Aided Medical Decision Making" Weiss, S., Kulikowski, C., Amarel, S., Safir, A., (1978) pp. 145-172, AI Journal, North Holland Press, 1978.

EXPERT--System for designing and applying consultation models using a relatively simple language to describe the models. It has been used in a variety of medical and non-medical applications (mainly rheumatology, ophthalmology and endocrinology).

Documentation available:

"A Guide to the Use of the EXPERT Consultation System" Weiss, S., Kern, K., and Kulikowski, C., CBM-TR-94, 11/78.

"EXPERT: A System for Developing Consultation Models," Proceedings IJCAI, Tokyo, August 1979, pp. 942-947, (also Rutgers Computer Science Report CBM-TR-97).

Anyone interested in using CASNET or EXPERT may contact either Kulikowski@Rutgers or Weiss@Rutgers.

REFERENCES

Amarel, S.: Computer-based interpretation and modeling in medicine and psychology: The Rutgers Research Resource. IN Siler and Lindberg (eds.), Computers in Life Science Research. Foseb and Planum, 1975.

Schmidt, C.F., Sridharan, N.S., and Goodson, J.L.: The plan recognition problem: An intersection of psychology and artificial intelligence. AI Journal 11(1,2), August, 1978 (special issue on applications to the sciences and medicine).

Weiss, S., Kulikowski, C.A., Amarel, S. and Safir, A.: A model-based method for computer-aided medical decision-making. AI Journal 11(1,2), August, 1978 (special issue on applications to the sciences and medicine).

National AIM Project: SIMULATION OF COGNITIVE PROCESSES (SCP)

Principal Investigators: James G. Greeno, Ph.D. (GREENO@SUMEX-AIM)
Alan M. Lesgold, Ph.D. (LESGOLD@SUMEX-AIM)
Learning Research and Development Center
University of Pittsburgh
Pittsburgh, Pennsylvania 15260
Dr. Lesgold: (412) 624-4901

The general purpose of the SCP Project is to develop increased understanding of normal and deficient cognitive functions, especially in reading and mathematics. Earlier work included simulations of interactive processes of grapheme-phoneme decoding and word recognition, and of semantic processes in comprehension of quantitative information in arithmetic word problems. The main emphasis at this time is on a collaboration with John Anderson, using the ACTF system to explore mechanisms of learning in the domain of geometry proofs. The SCP part of this work includes development of a system that learns by reading example proofs. The goal is to identify conceptual structures that are required for a learner to acquire planning strategies.

SOFTWARE AVAILABLE ON SUMEX

Programs are in a developmental stage and not yet available for use.

REFERENCES

Greeno, J.G., Magone, M.E. and Chaiklin, S.: Theory of constructions and set in problem solving. IN Memory and Cognition (In press). (Also available as Technical Report 1979/9, Learning Research and Development Center, Univ. Pittsburgh.)

Greeno, J.G.: Preliminary steps toward a cognitive model of learning primary mathematics. IN K. Fuson and W. Geeslin (eds.), Models of Children's Mathematical Learning, ERIC Information Center. (In press)

Lesgold, A.M. and Curtis, M.E.: Learning to read words efficiently. IN A. Lesgold and C. Perfetti (eds.), Interactive Processes in Reading, Erlbaum, Hillsdale, N.J. (In progress)

National AIM Project: SIMULATION AND EVALUATION OF CHEMICAL SYNTHESIS (SECS)

Principal Investigator: W. Todd Wipke, Ph.D.
Department of Chemistry
University of California at Santa Cruz
Santa Cruz, California 95064
(408) 429-2397 (WIPKE@SUMEX-AIM)

The SECS Project aims at developing practical computer programs to assist investigators in designing syntheses of complex organic molecules of biological interest. Key features of this research include the use of computer graphics to allow chemist and computer to work efficiently as a team, the development of knowledge bases of chemical reactions, and the formation of plans to reduce the search for solutions. SECS is being used by the pharmaceutical industry for designing syntheses of drugs.

A spin-off project, XENO, is aimed at predicting the plausible metabolites of foreign compounds for carcinogenicity studies. First, the metabolism is simulated; then the metabolites are evaluated for possible carcinogenicity.

SOFTWARE AVAILABLE ON SUMEX

SECS-- Available with a reaction library of over 400 reactions. The user needs a TTY or a DEC GT40 type graphics terminal.

XENO-- (for prediction of metabolites of xenobiotic compounds) is available for preliminary exploration since the project is still in the early development stages.

PRXBLD--(for building approximate molecular models from two-dimensional molecular models) is an energy minimization approach which is available both stand-alone and included within SECS.

REFERENCES

Spann, M.L., Chu, K.C., Wipke, W.T. and Ouchi, G.: Use of computerized methods to predict metabolic pathways and metabolites. J. Env. Pathology and Toxicology 2:123, 1978.

Wipke, W.T., Smith, G., Choplin, F. and Sieber, W.: SECS--Simulation and Evaluation of Chemical Synthesis: Strategy and planning. IN Computer-assisted Organic Synthesis Planning. ACS Symposium Series, 1977, pp. 97-127.

Wipke, W.T., Ouchi, G. and Krishnan, S.: Simulation and Evaluation of Chemical Synthesis--SECS. An application of artificial intelligence techniques. Artificial Intelligence 10:999, 1978.

National AIM Project: PROBLEM SOLVING EXPERTISE

Principal Investigators: Paul E. Johnson, Ph. D.
Center for Research in Human Learning
205 Elliott Hall
University of Minnesota
Minneapolis, Minnesota 55455
(612) 373-5302 (PJOHNSON@SUMEX-AIM)

William B. Thompson, Ph. D.
Department of Computer Science
136 Lind Hall
University of Minnesota
Minneapolis, Minnesota, 55455
(612) 373-0132 (THOMPSON@SUMEX-AIM)

The Minnesota SOLVER project focuses upon the development of strategies for discovering and representing the knowledge and skill of expert problem solvers. Although in the last fifteen years great progress has been made in synthesizing the expertise required for solving complex problems, most expert systems embody only the limited amount of expertise that individuals are able to report in a particular constrained language (e.g. production rules). What is still lacking is a theoretical framework capable of reducing dependence upon the expert's intuition or on the near exhaustive testing of possible organizations. Our methodology consists of: (1) extensive use of verbal thinking aloud protocols as a source of information from which to make verbal inferences about underlying cognitive structure and process; (2) development of computer models as a means of testing the adequacy of inferences derived from protocol studies; (3) testing and refinement of the cognitive models based upon the study of human and model performance in experimental settings. We are currently investigating problem solving expertise in domains of medicine, science, and law.

SOFTWARE AVAILABLE ON SUMEX

A version of the DIAGNOSER simulation model is being implemented on SUMEX; it should be available sometime this year.

REFERENCES

Johnson, P.E., Barreto, A., Hassebrock, F., Moller, J., Prietula, M., Feltovich, P., & Swanson, D. Expertise and error in diagnostic reasoning. Cognitive Science (in press). Johnson, P.E. Cognitive Models of Medical Problem Solvers. In D.C. Connelly, E. Benson, & D. Burke (eds.), Clinical decision making and laboratory use. University of Minnesota Press (in press). Johnson, P.E., & Thompson, W.B. Strolling down the garden path: Error prone tasks in expert problem solving. Proceedings of the Seventh International Joint Conference on Artificial Intelligence, Vancouver, British Columbia, August, 1981.

Stanford Project: GENERALIZATION OF AI TOOLS (AGE)

Principal Investigators: H. Penny Nii (NII@SUMEX-AIM)
Edward A. Feigenbaum, Ph.D.
Department of Computer Science
Stanford University
Stanford, California 94305
H.P. Nii: (415) 497-2739

The long-range objective of AGE, a SUMEX CORE RESEARCH Project, is to build a software laboratory for building knowledge-based, application programs. It is an attempt to define and accumulate knowledge-engineering tools, with rules to guide in the use of these tools. The design and implementation of the AGE program will be based primarily on the experiences gained in building knowledge-based programs by the Stanford Heuristic Programming Project in the last decade (The programs that have been or are being built are: DENDRAL, META-DENDRAL, MYCIN, HASP, AM, MOLGEN, GUIDON, CRYSALIS, PUFF, VM and SACON.). The initial AGE program contains a collection of tools suitable for constructing user programs based on the Blackboard paradigm (used in HASP and CRYSALIS) and Backward-chained production rules (used in MYCIN). In addition, AGE has facilities to aid the user in the construction, debugging, and running of his program.

SOFTWARE AVAILABLE ON SUMEX

AGE-1 is available on an experimental basis to a limited number of users. A public version of the programs, together with reference manuals and user guides, is planned for July, 1980.

REFERENCES

Nii, H.P. and Feigenbaum, E.A.: Rule-based understanding of signals. IN D.A. Waterman and F. Hayes-Roth (eds.), Pattern-directed Inference Systems. Academic Press, 1978, pp. 483-501.

Nii, H.P. and Aiello, N.: AGE: A knowledge-based program for building knowledge-based programs. Proc. Sixth IJCAI, Tokyo, August, 1979, pp. 645-655.

Aiello, N., Bock, C., Nii, H.P. and White, W.: Joy of AGE-ing: An Introduction to the Use of AGE-1 System, HPP Technical Memo, July, 1980.

Stanford Project: HANDBOOK OF ARTIFICIAL INTELLIGENCE

Principal Investigator: Edward A. Feigenbaum, Ph.D.
Avron Barr, Research Associate (BARR@SUMEX-AIM)
Department of Computer Science
Stanford University
Stanford, California 94305

The AI Handbook Project is a part of SUMEX CORE RESEARCH aimed at making the important results of AI research accessible to the large, multi-disciplinary community of scientists who want to build AI systems in their own problem areas. Students and researchers at Stanford and other AI laboratories have prepared over 300 short articles describing the fundamental ideas, useful techniques, and exemplary programs developed in the field over the last 20 years. These articles have been written for computer-literate scientists and engineers in other fields who are unfamiliar with AI research and jargon. The Handbook will provide a scientist who, for instance, might want to know what a "heuristic" is or how to build a "natural language" front end, with information about all of the relevant AI techniques and existing systems, as well as abundant pointers into the field's literature.

The Handbook is being published in report and book form. It also will be made available to the SUMEX community via an on-line information retrieval system. Following is a TOPIC OUTLINE for Volumes I and II: HANDBOOK OF ARTIFICIAL INTELLIGENCE

INTRODUCTION: The Handbook of Artificial Intelligence; Overview of AI Research; History of AI; An Introduction to the AI Literature

SEARCH: Overview; Problem Representation; Search Methods for State Spaces, AND/OR Graphs, and Game Trees; Six Important Search Programs

REPRESENTATION OF KNOWLEDGE: Issues and Problems in Representation Theory; Survey of Representation Techniques; Seven Important Representation Schemes;

AI PROGRAMMING LANGUAGES: Historical Overview of AI Programming Languages; Comparison of Data Structures and Control Mechanisms in AI Languages; LISP

NATURAL LANGUAGE UNDERSTANDING: Overview - History and Issues; Grammars; Parsing Techniques; Text Generation Systems; Machine Translation; The Early NL Systems; Six Important Natural Language Processing Systems

SPEECH UNDERSTANDING SYSTEMS: Overview - History and Design Issues; Seven Major Speech Understanding Projects

APPLICATIONS-ORIENTED AI RESEARCH--SCIENCE AND MATHEMATICS: Overview; TEIRESIAS - Issues in Expert Systems Design; Research on AI Applications in Mathematics (MACSYMA and AM); Research on AI Applications in Chemistry (DENDRAL, CRYSALIS, etc.); Other Scientific Applications Research

APPLICATIONS-ORIENTED AI RESEARCH--MEDICINE: Overview of Medical Applications Research; Six Important Medical Systems

APPLICATIONS-ORIENTED AI RESEARCH--EDUCATION: Historical Overview of AI Research in Educational Applications; Issues and Components of Intelligent CAI Systems; Seven Important ICAI Systems

AUTOMATIC PROGRAMMING: Overview; Techniques for Program Specification; Approaches to AP; Eight Important AP Systems

The following sections of the Handbook are still in preparation and will appear in Volume III: Theorem Proving; Vision; Robotics; Information Processing Psychology; Learning and Inductive Inference; Planning and Related Problem-solving Techniques.

Stanford Project: DENDRAL--RESOURCE RELATED RESEARCH - COMPUTERS IN CHEMISTRY

Principal Investigator: Carl Djerassi, Ph.D.
Department of Chemistry
Stanford University
Stanford, California 94305

Contact: Dr. Dennis SMITH@SUMEX-AIM
(415) 497-3144

The DENDRAL Project involves research in computer-assisted structure elucidation of unknown organic compounds of biological importance. This research has three major components: 1) program development; 2) biochemical applications; and 3) resource-sharing.

Recent program developments have been directed toward building more powerful interactive programs to assist chemists in the three major areas of structure elucidation: analysis of data to yield substructural information about an unknown ("planning"), advanced methods for assembly of substructures into complete structures ("structure generation"), and the prediction of data for structural candidates to rank-order the candidates by comparison of predicted and observed data ("testing"). Important problems of structure representation have been solved which have enabled dealing with stereochemical (three-dimensional) aspects of structure throughout the procedures.

Major areas of application of the programs in the research of this group and other collaborative projects include: a) marine natural products, particularly marine steroids and halogenated compounds which display biological activity; b) antibiotics and other derivatives of known or potential drugs; c) terpene alkaloids; d) photoproducts related to vitamin A; and e) conformational studies of narcotic analogs and polypeptides.

These programs are shared among a community of collaborators and guest users at SUMEX, with communication via computer network from a variety of sites in the U.S., Europe and Australia. Exportable versions of some programs are maintained. These versions have been installed successfully in more than 15 research laboratories throughout the world.

SOFTWARE AVAILABLE ON SUMEX

CONGEN--An interactive program for structure generation to yield candidate structures for an unknown based on inferred substructural components (exportable).

GENOA--An advanced structure generator capable of handling overlapping substructural information; uses CONGEN as a core component (exportable).

Meta-DENDRAL--An INTSUM, RULEGEN and RULEMOD sequence for automatic rule formation to relate observed data to substructures in mass spectrometry and carbon magnetic resonance spectroscopy.

REACT--A program for carrying out a complex sequence of chemical reactions and exploration of the consequences of those reactions.

NMR--For substructural inference and spectrum prediction in carbon magnetic resonance spectroscopy (will be exportable).

REFERENCES

Carhart, R.E., Varkony, T.H. and Smith, D.H.: Computer assistance for the structural chemist, IN D.H. Smith (ed.), Computer-Assisted Structure Elucidation. American Chemical Society, Washington, D.C., 1977, p. 126.

Djerassi, C., Smith, D.H. and Varkony, T.: A novel role of computers in the natural products field. Naturwiss. 66:9, 1979.

Nourse, J.G., Carhart, R.E., Smith, D.H. and Djerassi, C.: Exhaustive generation of stereoisomers for structure elucidation. J. Am. Chem. Soc. 101:1216, 1979.

Stanford Project: EXPEX -- THE EXPERT EXPLANATION PROJECT

Principal Investigator: Edward H. Shortliffe, M.D., Ph.D.
Departments of Medicine and Computer Science
Stanford University
Stanford, California 94305
(415) 497-6970 (SHORTLIFFE@SUMEX-AIM)

EXPEX is a new Stanford project that joined the AIM community in the Spring of 1981. The major thrust of the research is the development of new representation schemes to facilitate knowledge acquisition and explanation. This includes not only the study of fundamental representational formalisms but also the encoding of various types of knowledge, such as causal information and user models. The explanation portion of the research effort is dealing with a medical domain (endocrinology) and is being undertaken on SUMEX, whereas the knowledge acquisition portion deals with nonmedical topics and uses other computing resources at Stanford.

Our interest in explanation derives from the insights we gained in developing explanatory capabilities for the MYCIN system. In that system and its descendents, we were able to generate intelligible explanations by taking advantage of a rule-based representation scheme. The limitations of the justifications generated using MYCIN's explanation capabilities have become increasingly obvious, however, and have led to improved characterization of the kinds of explanation capabilities that must be developed if clinical consultation systems are to be accepted by physicians. EXPEX is devoted to the development of new theoretical insights into this problem.

REFERENCE

Wallis, J.W. and Shortliffe, E.H. Explanatory power for expert systems: studies in the representation of causal relationships for medical consultations. Internal technical report, Heuristic Programming Project, May 1981.

Stanford Project: MOLGEN--AN EXPERIMENT PLANNING SYSTEM FOR MOLECULAR GENETICS

Principal Investigators: Edward A. Feigenbaum, Ph.D.
Department of Computer Science
Stanford University

Laurence H. Kedes, M.D. (KEDES@SUMEX-AIM)
Department of Medicine
Stanford University
Stanford, California 94305
(415) 497-5897

Contact: Dr. Peter FRIEDLAND@SUMEX-AIM
(415) 497-1740

The goal of the MOLGEN Project is to apply the techniques of artificial intelligence to the domain of molecular biology with the aim of providing assistance to the experimental scientist. The most substantial problem under consideration is the task of experiment design. Two major approaches to this problem have been explored, one which instantiates abstracted experimental strategies with specific laboratory tools, and one which creates plans in toto, heavily influenced by the role played by interactions between plan steps. As part of the effort to build an experiment design system, a knowledge representation and acquisition package--the UNITS System, has been constructed. A large knowledge base, containing information about nucleic acid structures, laboratory techniques, and experiment-design strategies, has been developed using this tool. Smaller systems, such as programs which analyze primary sequence data for homologies and symmetries, have been built when needed.

SOFTWARE AVAILABLE ON SUMEX

Knowledge-based Experiment Design system (Friedland).

Meta-planning with Constraints experiment design system (Stefik).

UNITS system for knowledge representation and acquisition.

Interactive KORN Program for DNA sequence analysis.

GA1 program for restriction map construction.

SAFE program for gene excision.

REFERENCES

Friedland, P.E.: Knowledge-based experiment design in molecular genetics (Ph.D. thesis). Computer Science Dept. Report, CS-79-771, Stanford Univ.

Friedland, P.E.: Knowledge-based experiment design in molecular genetics. Proc. Sixth IJCAI, Tokyo, August, 1979, pp. 285-287.

Stefik, M.J.: An examination of a frame-structured representation system. Proc. Sixth IJCAI, Tokyo, August, 1979, pp. 845-852.

Stefik, M.J.: Planning with constraints (Ph.D. thesis). Computer Science Dept. Report, Stanford Univ. (In progress)

Stanford Project: The MYCIN PROJECTS -- KNOWLEDGE ENGINEERING FOR MEDICAL CONSULTATION

Principal Investigators: Bruce G. Buchanan, Ph.D.
Department of Computer Science
Stanford University
Stanford, California 94305
(415) 497-0935 (BUCHANAN@SUMEX-AIM)

Edward H. Shortliffe, M.D., Ph.D.
Departments of Medicine and Computer Science
Stanford University
Stanford, California 94305
(415) 497-6970 (SHORTLIFFE@SUMEX-AIM)

Subproject Directors:
MYCIN: Dr. Shortliffe, and A. Carlisle Scott
EMYCIN: Dr. Buchanan, and William van Melle
GUIDON: Dr. Buchanan, and William J. Clancey
ONCOCIN: Dr. Shortliffe, and A. Carlisle Scott

The MYCIN Projects are a set of research efforts overseen by a collaborative group of physicians and computer scientists who are developing intelligent systems using the techniques of knowledge engineering. The research focus includes knowledge acquisition, inexact reasoning, explanation, education, and the representation of time and of expert thinking patterns. Project members currently are working in a variety of medical domains including infectious disease therapy selection, intelligent computer-aided instruction, and the management of cancer chemotherapy protocols. Recent emphasis in the research has included intensive work regarding human engineering in an effort to implement the cancer therapy system for routine use by physicians. There is also a heightened interest in gearing representation, knowledge acquisition, and explanation more to the way that an expert actually thinks.

MYCIN Project (continued)

SOFTWARE AVAILABLE ON SUMEX

MYCIN--A consultation system designed to assist physicians with the selection of antimicrobial therapy for severe infections. It has achieved expert level performance in formal evaluations of its ability to select therapy for bacteremia and meningitis. Although MYCIN is no longer the subject of an active research program, the system continues to be available on SUMEX for demonstration purposes and as a testing environment for other research projects.

EMYCIN--The "essential MYCIN" system is a generalization of the MYCIN knowledge representation and control structure. It is designed to facilitate the development of new expert consultation systems for both clinical and non-medical domains.

GUIDON--A system developed for intelligent computer-aided instruction. Although it is being developed in the context of MYCIN's infectious disease knowledge base, the techniques are generalizable to any EMYCIN domain. The current research emphasis has been on an improved understanding of how the expert thinks so as to optimize the learning experience for the student.

ONCOCIN--This newest subproject is a system designed to provide oncologists with consultations regarding the management of patients receiving chemotherapy for cancer. Much of the knowledge in this domain is already well-specified in protocol documents, but expert judgments also need to be understood and modeled. Other research topics include human engineering and the development of capabilities to help achieve clinical acceptance of the program.

REFERENCES

Clancey, W.J., Shortliffe, E.H. and Buchanan, B.G.: Intelligent computer-aided instruction for medical diagnosis. Proc. Third Symposium on Computer Applications in Medical Care, Silver Spring, Maryland, October, 1979, pp. 175-183.

Shortliffe, E.H. Consultation systems for physicians. Proceedings of the CSCSI/SCEIO Conference, 14-16 May 1980, University of Victoria, British Columbia, pp. 1-11.

Shortliffe, E.H. The computer as clinical consultant (editorial). Arch. Int. Med., March 1980.

van Melle, W. A domain-independent system that aids in constructing knowledge-based consultation programs. PhD thesis, Computer Science Department, Stanford University, June 1980.

Stanford Project: PROTEIN STRUCTURE MODELING (CRYSALIS)

Principal Investigator: Edward A. Feigenbaum, Ph.D.
Department of Computer Science
Stanford University

Contact: Allan TERRY@SUMEX-AIM
(415) 497-1740

The CRYSALIS system is an application of artificial intelligence methodology to the task domain of protein crystallography. The focus is the structure determination problem: the derivation of an atomic model of the protein from an indistinct image of the electron density. The crystallographer interprets these data in light of the known chemical composition of the protein, general principles of protein chemistry, and his own experience. The goal of the CRYSALIS Project is to integrate these diverse sources of knowledge and data into a program that matches the crystallographer's level of performance in electron density map interpretation. A successful solution to this problem must deal with issues such as representation and management of a large knowledge base, opportunistic reasoning, and appropriate description of the emerging hypothesis, while keeping human engineering considerations in sight. Automation of this task would shorten the time for protein determination by several weeks to several months and would fill a major gap in the construction of a fully-automated system for protein crystallography.

SOFTWARE AVAILABLE ON SUMEX

CRYSTALLOGRAPHIC DATA REDUCTION PROGRAMS (in FORTRAN):
- A density map skeletonizer (SKEL37) based on an improved version of Greer's algorithm.
- A package for locating the critical points in a map.
- A general map-manipulation utility (INSPCT) that can find peaks, display regions, and compute various statistics.

TWO LISP SYSTEMS (with the caveat that both are under active development):
- A system (SEGLABELING) which heuristically parses the segmented map into labels similar to those a crystallographer would use.
- The inference system (CRYSALIS).

REFERENCES

Engelmore, R.S. and Nii, H.P.: A knowledge-based system for the interpretation of protein x-ray crystallographic data. Heuristic Programming Project Report HPP-77-2, Computer Science Dept., Stanford Univ., January, 1977.

Engelmore, R. and Terry, A.: Structure and function of the CRYSALIS system. Proc. Sixth IJCAI, Tokyo, August, 1979.

Nii, H.P. and Feigenbaum, E.A.: Rule-based understanding of signals. Heuristic Programming Project Report HPP-77-7, Computer Science Dept., Stanford Univ., April, 1977.